บทที่ 3

การดำเนินการวิจัย

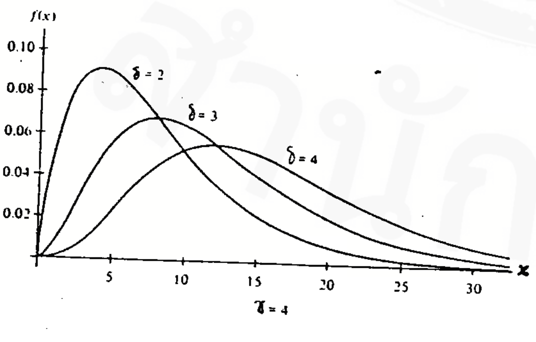
การดำเนินการวิจัยเพื่อหาข้อสรุปในการเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบความเท่ากันของค่าเฉลี่ยประชากร 2 กลุ่มที่เป็นอิสระกัน ในกรณีที่ค่าความแปรปรวนของประชากรทั้ง 2 กลุ่มแตกต่างกันและไม่ทราบค่า ได้แก่ สถิติทดสอบ Fitted Test สถิติทดสอบแมนวิทนีย์วิลคอกซันส่วนขยาย (Modified Mann-Whitney Wilcoxon : $\hat{U}$ ที่เสนอโดยไมเคิล เอ ฟลิกเนอร์ และ จอร์จ อี พอลิเซลโล, 1981) และสถิติทดสอบแมนวิทนีย์วิลคอกซันส่วนขยาย (Modified Mann-Whitney Wilcoxon : SU ที่เสนอโดย เอช เอช เลมเมอร์, 1987) โดยมีลักษณะข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงแบบปกติ คือ มีลักษณะการแจกแจงประชากรเป็นแบบเบ้ ได้แก่ การแจกแจงแกมมา และการแจกแจงไวบูลล์ โดยมีค่าความแปรปรวนของประชากรทั้งสองกลุ่มแตกต่างกันและไม่ทราบค่า มีขั้นตอนการดำเนินการดังนี้

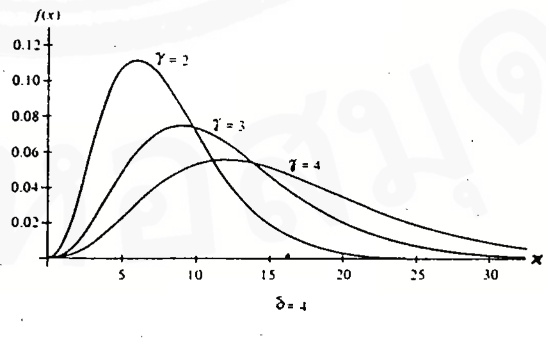
1. กำหนดการแจกแจงที่ศึกษา คือ การแจกแจงที่มีลักษณะแบบเบ้ ได้แก่

1.1 การแจกแจงแกมมา มีฟังก์ชันความหนาแน่น คือ

$$f(x;\delta,\gamma) = \frac{x^{(\delta-1)}\exp[-(\frac{x}{\gamma})]}{\Gamma(\delta)\gamma^{\delta}} \quad ; x \geq 0, \quad \delta > 0, \quad \gamma > 0$$

โดยมีค่าเฉลี่ยเท่ากับ $\delta\gamma$ และค่าความแปรปรวนเท่ากับ $\delta\gamma^2$



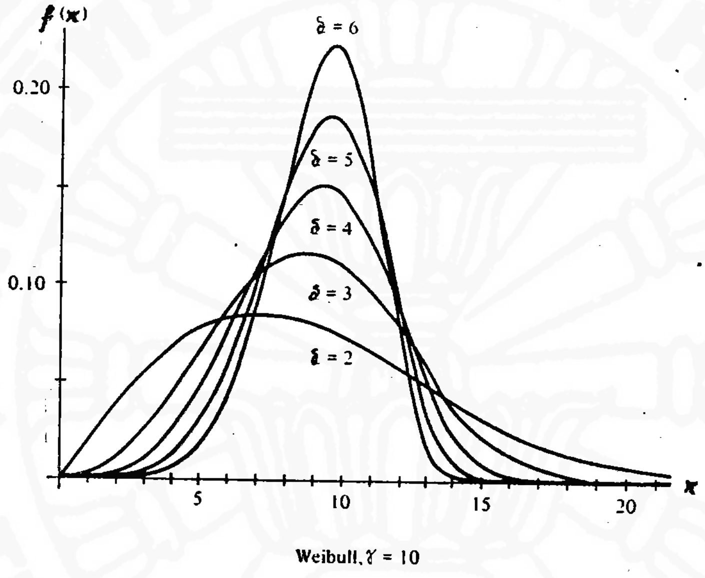


แสดงโค้งของการแจกแจงแกมมา ที่มา: Robert V. Hogg and Elliot A. Tanis, p.219

1.2 การแจกแจงไวบูลล์ มีฟังก์ชันความหนาแน่น คือ

$$f(x;\delta,\gamma) = \frac{\delta}{\gamma^{\delta}} x^{\delta-1} \exp[-(\frac{x}{\gamma})^{\delta}] \quad , \quad x \geq 0, \delta > 0, \gamma > 0$$

โดยมีค่าเฉลี่ยเท่ากับ $\gamma\Gamma(\frac{1}{\delta}+1)$ และค่าความแปรปรวนเท่ากับ $\gamma^2\{\Gamma(\frac{2}{\delta}+1)-[\Gamma(\frac{1}{\delta}+1)]^2\}$



แสดงโค้งของการแจกแจงไวบูลล์ ที่มา: Robert V. Hogg and Elliot A. Tanis, p.238

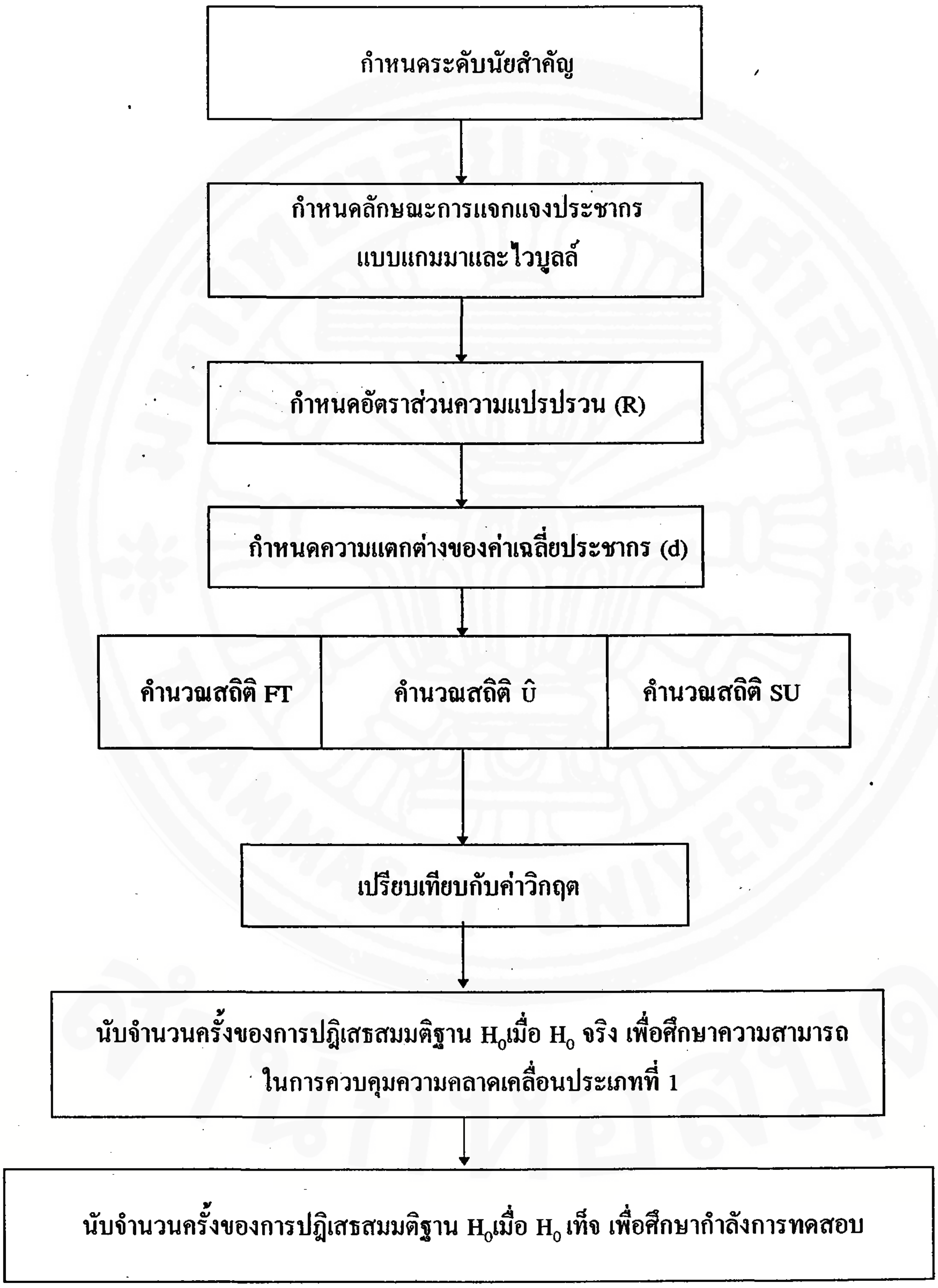
2. กำหนดขนาดตัวอย่าง คือ 6 และ10 โดยที่ขนาดตัวอย่างของประชากรกลุ่มที่ 1 มากกว่าขนาดตัวอย่างของประชากรกลุ่มที่ 2 ($n_1 \geq n_2$) แบ่งเป็น (n_1, n_2) ดังนี้ (6, 6), (10, 10), และ (10, 6)

3. กำหนดความแตกต่างของค่าเฉลี่ยประชากรทั้งสองกลุ่ม (d) คือ 0, 2, 5 และ 20

4. กำหนดสัดส่วนความแปรปรวนของประชากรกลุ่มที่ 1 : กลุ่มที่ 2 (R) ดังนี้ 1.5, 2, 3, และ 5

5. วิธีการดำเนินการทดลอง

5.1 ขั้นตอนการวิจัย



5.2 การจำลองข้อมูลตัวอย่างให้มีลักษณะตามที่กำหนดไว้ในขอบเขตการศึกษา โดยใช้โปรแกรมสำเร็จรูป MINITAB ด้วยคำสั่ง Random โดยมีขั้นตอนดังนี้

ระบุขนาดของข้อมูลที่ต้องการสร้าง โดยกำหนดจำนวนแถวและคอลัมน์ดังนี้

แถว(rows) เท่ากับจำนวนการทำซ้ำ (N)

คอลัมน์(columns) เท่ากับขนาดตัวอย่างข้อมูล(n)

ระบุชื่อการแจกแจงที่ต้องการพร้อมทั้งค่าพารามิเตอร์ของการแจกแจง แบ่งเป็นกรณีต่างๆ ดังนี้

<u>กรณีที่ 1</u> ให้ข้อมูลตัวอย่างมีการแจกแจงแกมมา, ผลต่างค่าเฉลี่ยประชากร(d) และสัดส่วนค่าความแปรปรวนประชากร(R)

<u>เงื่อนไขในการจำลองข้อมูล</u>

ประชากร 2 กลุ่มมีค่าเฉลี่ยแตกต่างกันเท่ากับ d นั่นคือ

$$\mu_1 - \mu_2 = d$$

$$\delta_1\gamma_1 - \delta_2\gamma_2 = d \qquad \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} 1$$

และสัดส่วนของความแปรปรวนประชากรในกลุ่มที่ 1 ต่อความแปรปรวนประชากรในกลุ่มที่ 2 เท่ากับ R นั่นคือ

$$\frac{\sigma_1^2}{\sigma_2^2} = R$$

$$\sigma_1^2 = R\sigma_2^2$$

$$\delta_1\gamma_1^{\;2} = R\delta_2\gamma_2^{\;2} \qquad \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} 2$$

<u>กำหนดค่า</u> $d = 0$, R และขนาดตัวอย่างตามต้องการ จากสมการที่ 1 และ 2

จะได้สมการในการคำนวณค่า พารามิเตอร์ของประชากรกลุ่มที่ 1 (δ_1, γ_1) และกลุ่มที่ 2 (δ_2 , γ_2) คือ

$$\delta_1 = \frac{\delta_2}{R}$$

$$\gamma_1 = R\gamma_2$$

กำหนดค่า d เป็นค่าใด ๆ ที่มากกว่า 0, R และขนาดตัวอย่างตามต้องการ พร้อมทั้งกำหนดค่าพารามิเตอร์ δ_1 และ กำหนดให้ $\gamma_1 = \gamma_2$ จากสมการที่ 1 และ 2 จะได้สมการในการคำนวณค่า พารามิเตอร์ของประชากรกลุ่มที่ 1 (γ_1)และกลุ่มที่ 2 (δ_2, γ_2) คือ

$$\delta_2 = \frac{\delta_1}{R}$$

$$\gamma_1 = \frac{d}{(1-\frac{1}{R})\delta_1}$$

ตัวอย่าง จำลองข้อมูลให้มีลักษณะการแจกแจงแกมมา, d=5, R=5 และขนาดตัวอย่างเท่ากัน (10,10) สุ่มข้อมูลซ้ำ 100 ครั้ง

กำหนดค่า $\delta_1 = 2$ จะได้ค่า $\gamma_1 = 3.125$, $\delta_2 = 0.4$ และ$\gamma_2 = 3.125$

โปรแกรมจำลองข้อมูลตัวอย่างกลุ่มที่ 1	MTB > random 100 c1-c10;
	SUB > gamma a=2 b=3.125.
โปรแกรมจำลองข้อมูลตัวอย่างกลุ่มที่ 2	MTB > random 100 c1-c10;
	SUB > gamma a=0.4 b=3.125.

กรณีที่ 2 ให้ข้อมูลตัวอย่างมีการแจกแจงไวบูลล์, ผลต่างค่าเฉลี่ยประชากร(d) และสัดส่วนค่าความแปรปรวนประชากร(R)

เงื่อนไขในการจำลองข้อมูล

ประชากร 2 กลุ่มมีค่าเฉลี่ยแตกต่างกันเท่ากับ d นั่นคือ

$$\mu_1 - \mu_2 = d$$

$$\delta_1\Gamma\left(\frac{1}{\delta_1}+1\right) - \delta_2\Gamma\left(\frac{1}{\delta_2}+1\right) = d \qquad \text{————— 3}$$

และสัดส่วนของความแปรปรวนประชากรในกลุ่มที่ 1 ต่อความแปรปรวนประชากรในกลุ่มที่ 2 เท่ากับ R นั่นคือ

$$\frac{\sigma_1^2}{\sigma_2^2} = R$$

$$\sigma_1^2 = R\sigma_2^2$$

$$\gamma_1^2\left[\Gamma\left(\frac{2}{\delta_1}+1\right)-\left(\Gamma\left(\frac{1}{\delta_1}+1\right)\right)^2\right] = R\gamma_2^2\left[\Gamma\left(\frac{2}{\delta_2}+1\right)-\left(\Gamma\left(\frac{1}{\delta_2}+1\right)\right)^2\right] \qquad \text{————— 4}$$

<u>กำหนดค่า</u> d, R และขนาดตัวอย่างตามต้องการ พร้อมทั้งกำหนดค่าพารามิเตอร์ δ_1 และ δ_2 คำนวณค่า พารามิเตอร์ของประชากรกลุ่มที่ 1 (γ_1)และกลุ่มที่ 2 (γ_2)

จากสมการ
$$\gamma_1 = \frac{d}{\left(1-\frac{1}{\sqrt{R}}\right)\Gamma\left(\frac{1}{\delta_1}+1\right)}$$

$$\gamma_2 = \frac{\gamma_1}{\sqrt{R}}$$

ในกรณีที่กำหนด ค่า $\delta_1 = 2$ สามารถคำนวณ γ_1 ได้จาก

$$\gamma_1 = \frac{2d\sqrt{R}}{(\sqrt{R}-1)\sqrt{\pi}}$$

<u>ตัวอย่าง</u> จำลองข้อมูลให้มีลักษณะการแจกแจงไวบูลล์, d=5, R=5 และขนาดตัวอย่างเท่ากัน (10,10) สุ่มข้อมูลซ้ำ 100 ครั้ง

กำหนดค่า $\delta_1 = 2$, $\delta_2 = 2$ จะได้ค่า $\gamma_1 = 10.20423$ และ $\gamma_2 = 4.56347$

โปรแกรมจำลองข้อมูลตัวอย่างกลุ่มที่ 1
```
MTB > random 100 c1-c10;
SUB > weibull a=2 b=10.20423.
```
โปรแกรมจำลองข้อมูลตัวอย่างกลุ่มที่ 2
```
MTB > random 100 c1-c10;
SUB > weibull a=2 b=4.56347.
```
(ตัวอย่างโปรแกรมการจำลองข้อมูล แสดงในภาคผนวก ค)

5.3 การทดสอบสมมติฐานเกี่ยวกับ ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เพื่อศึกษาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และทดสอบสมมติฐาน เพื่อเปรียบเทียบกำลังของการทดสอบ

เมื่อทำการจำลองข้อมูลตัวอย่างได้ตามลักษณะการแจกแจงประชากร, ค่าความแตกต่างค่าเฉลี่ยประชากร (d) และสัดส่วนของค่าความแปรปรวนของประชากรทั้ง 2 กลุ่ม (R) พร้อมทั้งกำหนดระดับนัยสำคัญ (α_0) ตามต้องการเรียบร้อยแล้ว

- นำตัวอย่างทั้ง 2 กลุ่มที่จำลองได้ มาคำนวณค่าสถิติทดสอบทั้ง 3 ชนิด เทียบกับค่าวิกฤตเพื่อศึกษาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยการตัดสินใจปฏิเสธสมมติฐานว่างเมื่อ

สมมติฐานว่างนั้นเป็นจริง และศึกษากำลังการทดสอบ โดยการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างนั้นเป็นเท็จ ดังรายละเอียดต่อไปนี้ นำข้อมูลตัวอย่างที่จำลองได้มาคำนวณค่าสถิติทดสอบฟิตเต็ด (FT) แล้วนำค่าที่ได้ไปเทียบกับค่าวิกฤตที่คำนวณได้ ตัดสินใจยอมรับหรือปฏิเสธสมมติฐานว่าง จากนั้นนำตัวอย่างชุดใหม่มาทำการทดสอบ ทำเช่นนี้จนครบ 100 ครั้ง นับจำนวนการปฏิเสธสมมติฐานว่างทั้งหมด หลังจากนั้นก็นำตัวอย่างชุดเดียวกัน มาคำนวณค่าสถิติทดสอบแมนวิทนีย์วิลคอกซันส่วนขยาย (SU) เปรียบเทียบค่าสถิติที่คำนวณได้กับค่าวิกฤตของสถิติทดสอบ SU ทำเช่นนี้จนครบ 100 ครั้ง นับจำนวนการปฏิเสธสมมติฐานว่าง หลังจากนั้น ก็นำตัวอย่างชุดเดียวกันนี้ มาคำนวณค่าสถิติทดสอบแมนวิทนีย์วิลคอกซันส่วนขยาย ($\hat{U}$) เปรียบเทียบค่าสถิติที่คำนวณได้กับค่าวิกฤตของสถิติทดสอบ $\hat{U}$ ทำเช่นนี้จนครบ 100 ครั้ง นับจำนวนการปฏิเสธสมมติฐานว่าง จากนั้นก็ทำการเปลี่ยนรูปแบบการแจกแจง ขนาดตัวอย่าง และสัดส่วนค่าความแปรปรวนประชากร จนครบทุกกรณีศึกษา

5.3.1 การทดสอบสมมติฐานเกี่ยวกับ ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จะใช้สถิติทดสอบ Z และ การทดสอบสมมติฐานแบบสองทาง เพื่อศึกษาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

กำหนดให้ α แทน ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

$\alpha.$ แทน ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง

α_0 แทน ระดับนัยสำคัญที่กำหนด

N แทน จำนวนครั้งที่ทำการทดลอง

สมมติฐานที่ต้องการทดสอบ

$$Ho:\ \alpha = \alpha_0$$

$$H1:\ \alpha \neq \alpha_0$$

สถิติทดสอบ

$$Z = \frac{\alpha. - \alpha_0}{\sqrt{\alpha_0(1-\alpha_0)/N}}$$

กำหนดค่าระดับนัยสำคัญของการทดสอบ Z เท่ากับ 0.05 ดังนั้นสถิติทดสอบที่ใช้ในการศึกษาจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ไว้ได้ ก็ต่อเมื่อ

$$Z_{\frac{0.05}{2}} \leq Z \leq Z_{1-\frac{0.05}{2}}$$

จะได้เกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 เมื่อกำหนดระดับนัยสำคัญของการทดสอบ Z เป็น 0.05 และจำนวนครั้งที่ทำการทดลอง (N) เท่ากับ 100 ครั้ง สถิติทดสอบที่ศึกษาสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ไว้ได้ถ้า $-1.96 \leq Z \leq 1.96$ ซึ่งสามารถสรุปผลได้ 2 กรณี คือ

- กรณีที่ระดับนัยสำคัญของการศึกษา (α_0) เท่ากับ 0.05

$$Z_{0.025} \leq Z \leq Z_{0.975}$$

$$-1.96 \leq Z \leq 1.96$$

$$-1.96 \leq \frac{\alpha_{*} - .05}{\sqrt{.05(1-.05)/100}} \leq 1.96$$

$$0.0073 \leq \alpha_{*} \leq 0.092$$

ดังนั้น เมื่อกำหนดระดับนัยสำคัญของการศึกษา(α_0) เท่ากับ 0.05 สถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ไว้ได้ถ้า α_{*} อยู่ในช่วง (0.0073, 0.0927)

- กรณีที่ระดับนัยสำคัญของการศึกษา (α_0) เท่ากับ 0.01

$$Z_{0.025} \leq Z \leq Z_{0.975}$$

$$-1.96 \leq Z \leq 1.96$$

$$-1.96 \leq \frac{\alpha_{*} - .01}{\sqrt{.01(1-.01)/100}} \leq 1.96$$

$$0.0 \leq \alpha_{*} \leq 0.0295$$

ดังนั้น เมื่อกำหนดระดับนัยสำคัญของการศึกษา(α_0) เท่ากับ 0.01 สถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ไว้ได้ถ้า α_{*} อยู่ในช่วง (0.0, 0.0295)

5.3.2 การทดสอบสมมติฐานเกี่ยวกับกำลังของการทดสอบ

เปรียบเทียบค่ากำลังของการทดสอบ โดยใช้สถิติทดสอบ Z โดยทดสอบสมมติฐานแบบทางเดียวด้านขวา

กำหนด p_i แทน ค่ากำลังของการทดสอบของสถิติทดสอบที่ i

p_j แทน ค่ากำลังของการทดสอบของสถิติทดสอบที่ j

$\hat{p}_i$ แทน ค่ากำลังของการทดสอบจากการทดลองของสถิติทดสอบที่ i

$\hat{p}_j$ แทน ค่ากำลังของการทดสอบจากการทดลองของสถิติทดสอบที่ j

(ค่า $\hat{p}$ เท่ากับจำนวนครั้งของการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นเท็จหารด้วยจำนวนครั้งที่ทำการทดลองทั้งหมด)

N_i แทน จำนวนครั้งที่ทำการทดลองของสถิติทดสอบชนิดที่ i

N_j แทน จำนวนครั้งที่ทำการทดลองของสถิติทดสอบชนิดที่ j

สมมติฐานที่ต้องการทดสอบ

$Ho : p_i = p_j$

$H1 : p_i > p_j \quad ; i < j$

สถิติทดสอบ
$$Z = \frac{(\hat{p}_i - \hat{p}_j)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{N_i} + \frac{1}{N_j}\right)}}$$

โดยที่
$$\bar{p} = \frac{(N_i\hat{p}_i + N_j\hat{p}_j)}{(N_i + N_j)}$$

เมื่อ กำหนดระดับนัยสำคัญของการทดสอบ Z เท่ากับ 0.05 จะสามารถสรุปผลได้ดังนี้

- ถ้า Z ที่คำนวณได้ มีค่ามากกว่า 1.645 ที่ได้จากตารางการแจกแจงปกติมาตรฐานที่ระดับนัยสำคัญที่กำหนด จะสรุปว่า สถิติทดสอบที่ i มีกำลังการทดสอบมากกว่าสถิติทดสอบที่ j ที่ระดับนัยสำคัญ 0.05 (*)

- ถ้า Z ที่คำนวณได้ มีค่าน้อยกว่า 1.645 ที่ได้จากตารางการแจกแจงปกติมาตรฐานที่ระดับนัยสำคัญที่กำหนด จะสรุปว่า สถิติทดสอบที่ i และสถิติทดสอบที่ j มีกำลังการทดสอบไม่แตกต่างกัน ที่ระดับนัยสำคัญ 0.05 (*)

เมื่อ กำหนดระดับนัยสำคัญของการทดสอบ Z เท่ากับ 0.01 จะสรุปผลได้ดังนี้

- ถ้า Z ที่คำนวณได้ มีค่ามากกว่า 2.33 ที่ได้จากตารางการแจกแจงปกติมาตรฐานที่ระดับนัยสำคัญที่กำหนด จะสรุปว่า สถิติทดสอบที่ i มีกำลังการทดสอบมากกว่าสถิติทดสอบที่ j กัน ที่ระดับนัยสำคัญ 0.01 (**)

- ถ้า Z ที่คำนวณได้ มีค่าน้อยกว่า 2.33 ที่ได้จากตารางการแจกแจงปกติมาตรฐานที่ระดับนัยสำคัญที่กำหนด จะสรุปว่า สถิติทดสอบที่ i และสถิติทดสอบที่ j มีกำลังการทดสอบไม่แตกต่างกัน กัน ที่ระดับนัยสำคัญ 0.01 (**)

แผนภาพแสดงขั้นตอนการคำนวณสถิติ Fitted Test

ข้อมูลตัวอย่าง X_{1i}, X_{2j}

↓

คำนวณค่าเฉลี่ย ($\bar{x}_{\cdot}$) และค่าความแปรปรวน (S^2)
ของข้อมูลแต่ละกลุ่ม

↓

คำนวณค่าสถิติ $v = \dfrac{\bar{x}_1 - \bar{x}_2}{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^{\frac{1}{2}}}$

↓

คำนวณค่าวิกฤติ $V(\hat{C})$

↓

ทดสอบสมมติฐาน

แผนภาพแสดงขั้นตอนการคำนวณสถิติ $\hat{U}$

ข้อมูลตัวอย่าง X_{1i}, X_{2j}

↓

เรียงลำดับข้อมูลตัวอย่างร่วมกัน
ทั้งสองกลุ่มจากน้อยไปมาก

↓

ให้ค่าอันดับแก่ข้อมูลกลุ่มที่ 1 (Q_i)
และกลุ่มที่ 2 (R_j)

↓

คำนวณค่า P_i และ S_j

$P_i = Q_i - i$, $i = 1, ..., n_i$

$S_j = R_j - j$, $j = 1, ..., n_j$

↓

คำนวณค่าสถิติ $$\hat{U} = \frac{\left(\sum_{i=1}^{n_1} P_i - \sum_{j=1}^{n_2} S_j\right)}{2\sqrt{\sum_{i=1}^{n_1}(P_i - \bar{P})^2 + \sum_{j=1}^{n_2}(S_j - \bar{S})^2 + \overline{PS}}}$$

↓

ทดสอบสมมติฐาน

แผนภาพแสดงขั้นตอนการคำนวณสถิติ SU

ข้อมูลตัวอย่าง X_{1i}, X_{2j}

↓

หาค่าผลต่าง (X_{1i} - X_{2j})

↓

เรียงลำดับผลต่างจากน้อยไปมากแล้ว
(ไม่คำนึงถึงเครื่องหมาย)
ให้ค่าอันดับแก่ผลต่าง $r\left(\left|x_{1i}-x_{2j}\right|\right)$

↓

คำนวณสถิติ $SU=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} r\left(\left|x_{1i}-x_{2j}\right|\right)U_{ij}$

↓

ทดสอบสมมติฐาน